

## TELEPHONE LISTING

## (i) GENERAL INFORMATION:

- (i) APPLICANT: Iris Becker, Israel Vlodavsky and Elena  
Feldstein
- (ii) TITLE & INVENTION: POLYNUCLEOTIDE ENCODING A POLYPEPTIDE  
HAVING HEPATOPROTECTIVE ACTIVITY AND EXPRESSION  
OF SAME IN GENETICALLY MODIFIED CELLS
- (iii) NUMBER OF TELEPHONE: 47
- (iv) CORRESPONDENCE ADDRESS:
- (A) ADDRESSEE: Mark M. Friedman c/o Anthony Castorina  
(B) STREET: Joel Jefferson Davis Highway, Suite 207  
(C) CITY: Arlington  
(D) STATE: Virginia  
(E) COUNTRY: United States of America  
(F) ZIP: 22202
- (v) COMPUTED FEASIBLE FORM:
- (A) MEDIUM TYPE: 1.44 megabyte, 3.5" microdisk  
(B) COMPUTER: Twinhead\* Slimnote-890TE  
(C) OPERATING SYSTEM: MS DOS version 6.2,  
Windows version 3.11  
(D) SOFTWARE: Word for Windows version 2.0 converted to  
an ASCII file
- (vi) CURRENT APPLICATION DATA:
- (A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
- (A) APPLICATION NUMBER: 09/922,170  
(B) FILING DATE: 2 SEP 1997  
(A) APPLICATION NUMBER: 09/109,386  
(B) FILING DATE: 10 JUL 1998  
(A) APPLICATION NUMBER: PCT/US98/17954  
(B) FILING DATE: 31 AUG 1998
- (viii) ATTORNEY/AGENT INFORMATION:
- (A) NAME: Friedman, Mark M.  
(B) REGISTRATION NUMBER: 33,883  
(C) REFERENCE/DOCKET NUMBER: 910/14
- (ix) TELECOMMUNICATION INFORMATION:
- (A) TELEPHONE: 972-3-5625553  
(B) TELEFAX: 972-3-5625554  
(C) TELEX:

## INFORMATION FOR FILING IN THE:

- (i) FILING OFFICE:
- (A) NUMBER:  
(B) TIME: 10:00-12:00  
(C) TRANSMISSION: single  
(D) TIME: 10:00-12:00
- (ii) TELEPHONE LISTING IN THE:
- (A) TELEPHONE LISTING IN THE:

CGATGATGCGATGAAATGAAATG 14

(1) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 29  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ACTGACTATA GCGTCTGAGG CGG 29

(1) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CGATCTTAGG CGCTTTTCTT CG 22

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 15  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:3:

TTTTTTTTTT TTTT 15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 23  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTGGATGCGA AGAAGGAATG AAC 23

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGATGATGCGATGAAATGAAATG 14

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 5

1A	LENGTH:	1701
1B	TYPE:	1001-1 1001
1C	STANDARD NUMBER:	1001-1
1D	TEST NAME:	1001-1

[illegible]

## A. B. G. J. VAN DER KAM, J. H. VAN DIJK, and J. J. VAN DER VEGT

M + L<sub>1</sub> L<sub>2</sub> A<sub>1</sub> C<sub>1</sub> 177 11    A<sub>1</sub> L<sub>1</sub> 11    11    11    L<sub>1</sub> M + L<sub>1</sub> L<sub>2</sub>

Don Don City Pro Don City Fr Don Don Fr City Ala Don Fr Ala Pro

65	70	75	80
Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Phe Gly Gly			
85	90	95	100
Thr Lys Thr Asp Phe Leu Ile Phe Asp Ile Lys Lys Gln Ser Thr Phe			
105	110	115	120
Glu Glu Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asp Ile Cys Lys			
125	130	135	140
Tyr Gly Ser Ile Pro Pro Asp Val Glu Gln Lys Leu Arg Leu Glu Trp			
145	150	155	160
Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe			
165	170	175	180
Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe			
185	190	195	200
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu			
205	210	215	220
Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu			
225	230	235	240
Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe Ile Asn Gly Ser			
245	250	255	260
Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys Leu Leu Arg Lys Ser			
265	270	275	280
Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro Asp Val Gly Gln Pro Arg			
285	290	295	300
Arg Lys Thr Ala Lys Met Leu Lys Ser Phe Leu Lys Ala Gly Gly Gln			

Val Ile Asp Ser Val Thr Trp His Pro Thr Tyr Leu Arg Gly Arg Thr
305

Ala Thr Arg Glu Asp Phe Leu Asn Pro Asp Val Leu Asp Ile Phe Ile
315

Ser Ser Val Glu Lys Val Phe Gln Val Val Gln Ser Thr Arg Ile Gly
325

Val Thr Ile Ser Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr Thr
335

Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp His Asn His Asp His  
 434 437 440 443  
 Leu His Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
 444 447 450 453  
 Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Asn Asn Lys Leu Asn  
 454 457 460 463  
 Val Tyr Leu His Cys Thr Asn Thr Asp Asn Phe Asn Tyr Lys Gln Gly  
 464 467 470 473  
 Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr Lys Tyr Leu  
 474 477 480 483  
 Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
 484 487 490 493  
 Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys Ser Val Gln Leu Asn  
 494 497 500 503  
 Gly Leu Thr Leu Lys Met Val Asp Asp Gln Thr Leu Pro Pro Leu Met  
 504 507 510 513  
 Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser  
 514 517 520 523  
 Tyr Ser Phe Phe Val Ile Arg Asn Ala Lys Val Ala Ala Cys Ile  
 524 527 530 533

2. INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1721  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(ii) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ATG GTC CTA TGG TGA AA: CTT GCG CTC GAG TGG TGG CTA ATG CTC CTC 116  
 Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro Phe Leu Met Leu Leu  
 117 120 123  
 TTT CTC GAG GAG GTC GAG GAG GTC GAG GAG GTC GAG GAG GTC GAG GAG 119  
 Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro Gly Ala Leu His And Pro

Leu His Leu Val Ser His Ser Phe Leu Ser Val Thr Ile Asn Ala Asn  
 65 70 75 80

TTG GGC AAG GAG GCG GCG TTT CTC ATC CTC CTC GAT TGT CAA AAG TTT 85  
 Leu Ala Thr Asp Pro Arg Ile Leu Ile Leu Leu Gly Ser His Lys Leu  
 65 70 75 80

CGT ACC TTG GGC AAG GCG TTG TCT CCT GCG TAC CTC AAG TTT GGT GGC 90  
 Arg Thr Leu Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Ile Gly Gly  
 85 90 95

AGC AAG GCA GAG TTT CTA ATT TTC GAT GCG AAG AAG GAA TCA ACC TTT 95  
 Thr Lys Thr Asp Ile Leu Ile Phe Asp Pro Lys Lys Glu Ser Thr Phe  
 100 105 110

GAA GAG AAG AGT TAC TGG CAA TTT CAA CTC AAC CAG GAT ATT TTT AAA 100  
 Glu Ala Arg Ser Tyr Trp Gln Ser Gln Val Asn Gln Asn Ile Cys Lys  
 115 120 125

TAT GGA TCC ATC CCT CCT GAT GTG GAG GAG AAG TTA GCG TTG GAA TGG 105  
 Tyr Gly Ser Ile Pro Ile Asp Val Gln Glu Lys Leu Arg Leu Glu Trp  
 130 135 140

CCC TAC CAG GAG CAA TTG CTA CTC CGA GAA CAC TAC CAG AAA AAG TTC 110  
 Pro Tyr Gln Glu Gln Leu Leu Leu Arg Glu His Tyr Gln Lys Lys Phe  
 145 150 155 160

AAG AAC AGC ACC TAC TCA AGA AGC TCT GTA GAT GTG CTA TAC ACT TTT 115  
 Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe  
 165 170 175

GCA AAC TGC TCA CGA CTG GAC TTG ATC TTT GCG CTA AAT GCG TTA TTA 120  
 Ala Asn Lys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu Leu  
 180 185 190

AGA ACA GCA GAT TTG CAG TGG AAC AGT TTT AAT GCT CAG TTG CTC CTG 125  
 Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu Leu Leu  
 195 200 205

GAT TAC TTT TTT TTT AAA GCG TAT AAT AAT TTT TAA GAA CTA TTT AAT 130  
 Asp Tyr Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe Phe  
 210 215 220

GAA TTT AAT AAT TTT TTT AAT AAT TTT TTT AAT AAT TTT AAT AAT TTT 135  
 Gln His Asn Ser His Leu Lys Lys Ala Asn Ile Phe His Asn Gly Ser  
 225 230 235 240

TAT TTA GAA GAA GAT TAT ATT TAA TTT TAT AAA TTT CTA AAG AAT TTT 140  
 His Leu His His Arg Thr Ile His Leu His Leu His Leu His Leu  
 245 250 255

ATG TTT GAA GAA GAT TAT ATT TAA TTT TAT AAA TTT CTA AAG AAT TTT 145  
 Met Phe Gln Gln GAT TAT ATT TAA TTT TAT AAA TTT CTA AAG AAT TTT  
 150 155 160

GTG ATT GAT TGA GTT ACA TGG CAT CAG TAC TAT TTG AAT GGA CCG ACT 974  
Val Ile Asp Ser Val Thr Trp His His Tyr Tyr Leu Asn Gly Arg Thr

GTT ACC AAG GAA GAT TTT TTA AAA TTT GAT GTA TTG AAA ATT TTT ATT 1041  
Ala Thr Asn Gln Asp Phe Leu Asn Leu Asp Val Leu Asn Ile Phe Ile  
305 310 315 320

TGA TGT CTG CAA AAA GTT TTG CAG GTG GTT GAG AGC AGC AGG CTT GGT 1070  
Ser Ser Val Gln Lys Val Phe Gln Val Val Glu Ser Thr Arg Pro Gly  
325 330 335

AAG AAG CTC TGG TTA GGA GAA ACA AGC TGT GCA TAT CAA GGC GGA GCG 1118  
Lys Lys Val Trp Leu Gly Gln Thr Ser Ser Ala Tyr Gly Gly Gly Ala  
340 345 350

CCG TTG CTA TCC GAC ACC TTT GCA GGT GGC TTT ATG TGG CTG GAT AAA 1166  
Pro Leu Leu Ser Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys  
355 360 365

TTG GGC CTG TCA GCC CGA ATG GGA ATA GAA GTG GTG ATG AGG CAA GTA 1214  
Leu Gly Leu Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Gln Val  
370 375 380

TTG TTT GGA GCA GGA AAC TAC CAG TTA GTG GAT GAA AAC TTC GAT CCT 1262  
Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe Asp Pro  
385 390 395 400

TTA CCG GAT TAT TGG CTA TGT CTT CTG TTC AAG AAA TTG GTG GGC ACC 1310  
Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
405 410 415

AAG GTG TTA ATG GCA AGC GTG CAA GGT TCA AAG AGA AGG AAG CTT CGA 1358  
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu Arg  
420 425 430

GTA TAC CTT CAT TGC ACA AAC ACT GAC AAT CCA AGG TAT AAA GAA GGA 1406  
Val Tyr Leu His Tyr Thr Asn Thr Asp Asn Phe Arg Tyr Lys Gln Gly

AGT TTA GAT TAT TAT TTA AAA TTT GAT GTA TTG AAA ATT TTT ATT 1454  
Arg Leu Thr Tyr Tyr Ala Ile Asn Leu His Asn Val Thr Trp Tyr Leu  
435 440 445

TGG TTA CTT TAT TGT TTT TAT AAG AAG GAA GTG GAT AAA TAT CTT GTA 1502  
Arg Leu Phe Tyr Phe Phe Ser Asn Lys Gln Val Asp Lys Tyr Leu Leu  
445 450 455 460

1550 1555 1560 1565 1570 1575 1580 1585 1590 1595

GAA AAA GGT GTG GCG GCA GAG AGT TAA TTA TTT TTG CAA GGT TGC TTA 1-4  
Glu Lys Pro Leu Arg Pro Gly Ser Ser Leu Gly Leu Pro Ala Phe Ser

TAT AAT TTT TTT TGG ATA AAG GAT GGT AAA GTT GGT GGT TTT ATT TAA TGA TGG  
Tyr Gln Phe Phe Val Ile Ala Asn Ala Lys Val Ala Ala Gly Ile

530 535 540 543

[illegible]

## (2) INFORMATION FOR SEQ. II. 11.12:

## (1) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 734  
(B) TYPE: nucleic acid  
(C) STRAIN ENERGY: 4.46 eV  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12

CTGCGAAGAA	GCTCTGGTTG	GGAGAGATGA	GCTCAGTTTA	GGTGGGCGGT	GCACCCCTTC	60
TGTCCAAAC	CTTTGAGGT	GGCTTTATGT	GGCTGGATAA	ATTGGGCGTG	TCACGCCAGA	120
TGGGCATAGA	AGTCGPGATG	AGGCAGTGT	TCTTCGAGC	AGGCAACTAC	CACCTTAGTG	180
ATGAAAACTT	TGAGCCTTTA	CCTGATTAAT	GGCTCTTTCT	TCTGTTCAAG	AAACTGGTAG	240
GTCCACGGT	GTTACTGTCA	AGAGTGAAG	GCCAGAGAG	GAGCAAACTC	CGAGTGTATC	300
TCCACTGCAC	TAACTGTAT	CAGCAAGAT	ATCAGGAAGC	AGATCTAAT	CTGTATGTCC	360
TGAACCTCCA	TAATGTCACC	AAGCAGCTGA	AGGTACCCCT	TCCGTTGTTC	AGGAAACCAG	420
TGGATAGTA	CCTTCTGAAG	CCTTCGAGC	CGGATGAATT	ACTTTCCAAA	TCTGTCCAAC	480
TGAACGCTCA	AATTCTGAAG	ATGCTGATG	AGCAGAGCT	GCCAGCTTTG	ACAGAAAAAC	540
CTTTTCGCG	AGGAAGTGA	CTAAGCTTG	CTGCTTTTCT	CTATGGTTTT	TTTGTGATAA	600
GAAATGCCAA	AATCGCTGCT	TGTATATGA	AATAAAAGC	ATACGGTACC	CCTGAGACAA	660
AAGGCGAGGG	GGGTGTTATT	CATAAAAGAA	AACCTAGTT	TAGAGGGCCA	CCTCCTTGCC	720
GAGTTACAGA	GCTTCGGGAG	GGTGGGATAC	ACTTCAGTAT	TACATTCAAT	GCTGTGTTCT	780
CTGTAGGAAG	AATATTTCAG	CTGTGTGAGC	TTAATAGTAC	TGTC		825

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A)	LENGTH:	1869
(B)	TYPE:	nucleic acid
(C)	STRAN LENGTH:	4 base
(D)	TOTAL:	1 base

[illegible][illegible]



GAACTAGGCA ATGAAGCTAA CACTTTTATT AAGAACTTG ATATTTTAT CAATGGGTGG 900  
 CAGTTAGGAG AAGATTATAT TAAATTAT AAGCTTAA GAAATGAC CTTCAAAAAA 1020  
 GCAAAATCT ATCTCTGA TTTTGGTAG CTTGAAAAA AGAGGCTAA GATGCTGAAG 1080  
 AGTTCTCTGA AGCTCTGG AGAAGTATT GATTAGTTA CATGGCATCA CTACTATTG 1140  
 AATGGAGGSA CTCTACCAAG AGAAGATTT CTAATTTG ATCTATTGA CATTCTTATT 1200  
 TCACTCTCTC AAAAGCTTT TAACTCTTT GAGAGAGTA CGCTGGGAA GAACTCTCTG 1260  
 TTAGGAGAAA CAAGCTCTG ATATGGAGG GAGAGTCT TGCTATCGA CAGCTTTGCA 1320  
 CCTGGTTTA TGTGGGTGA AAAATCTGG CTGTAGGAG GAATGGGAAT AGAAGTGGTG 1380  
 ATAGGGAAG TATTCTTTG AAGAGTAA TAACTTTAG TGCTGAAAA CTTGATCCT 1440  
 TTCTCTGATT ATGGGTATC TTTCTTTT AAGAACTTG TGCTGAGAA GGTGTTAATG 1500  
 CCAAGCTCTC AAGGTCTAAA GAGAGAG AAG CTTGAGTAT ACCTTCATTG CACAAACAAT 1560  
 CAAATCTCA CCTATAAGAG AGAAGATTT ACTCTCTAG CATAAAGCT CATAAGCTG 1620  
 ACAAGTACT TGGGTTTAC CTATCTTTT TCTAAAGAG AAGTGGATAA ATACTCTCTA 1680  
 AGACCTTTG GACCTCATGG ATACTTTCT AAATCTCTG AACTCAATGG TCTAACTCTA 1740  
 AAGATGCTG ATGATGAAA CTTGCACTT TTAATGAAA AACCTCTCG GCCAGGAAGT 1800  
 TCACTGCTG TGGCAGCTT CTCTAATAG TTTTCTCTA TAAAGAAATG CAAAGTCTG 1860  
 CCTTGCATCT GAAATAAAAA TATACTACT CTGACACTG 1899

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14

Met Glu Gly Ala Val Gly Gly Val Arg Arg Arg Asn Gly Ala Glu  
 5 10 15  
 Glu Arg Arg Lys Gly Arg Trp Gly Ser Ala Gly Gly Ser Ala Arg  
 20 25 30  
 Ala Leu Asp Ser Pro Leu Arg Gly Ser Trp Arg Gly Glu Gln Pro  
 35 40 45  
 Gly Glu Pro Lys Met Leu Leu Arg Ser Lys Pro Ala Leu Pro Pro  
 50 55 60  
 Pro Leu Met Leu Leu Leu Gly Pro Leu Gly Pro Leu Ser Pro  
 65 70 75  
 Gly Ala Leu Pro Arg Pro Ala Gln Ala Gln Asp Val Val Asp Leu  
 80 85 90  
 Asp Phe Phe Thr Gln Glu Pro Leu His Leu Val Ser Pro Ser Phe  
 95 100 105  
 Leu Ser Val Thr Ile Asp Ala Ala Leu Ala Thr Arg Trp Arg Phe  
 110 115 120  
 Leu Ile Leu Leu Gly Ser Leu Lys Leu Arg Thr Leu Ala Arg Gly  
 125 130 135  
 Leu Ser Phe Ala Tyr Leu Arg Phe Gly Gly Thr Lys Thr Asp Phe  
 140 145 150  
 Leu Ile Phe Asp Phe Lys Lys Gln Ser Thr Phe Glu Glu Arg Ser  
 155 160 165  
 Tyr Trp His Ser Gln Val Asn Gln Arg Ile Lys Lys Trp Glu Thr  
 170 175 180

Leu Ala Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu	345	345
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu	346	346
Leu Gly Asn Gln Pro Asn Ser Ile Leu Lys Lys Ala Asp Ile Ile	347	347
Ile Asn Gly Ser Gln Leu Gly Gln Asp Tyr Ile Gln Leu His Lys	348	348
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro	349	349
Asp Val Gly Gln Pro Arg Ala Lys Thr Ala Lys Met Leu Lys Ser	350	350
Phe Leu Lys Ala Gly Gly Gln Val Ile Asp Ser Val Thr Trp His	351	351
His Tyr Tyr Leu Asn Gly Arg Thr Ala Thr Arg Gln Asp Phe Leu	352	352
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val	353	353
Phe Gln Val Val Gln Ser Thr Arg Pro Gly Lys Lys Val Trp Leu	354	354
Gly Gln Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro Leu Leu Ser	355	355
Asp Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys Leu Gly Leu	356	356
Ser Ala Arg Met Gly Ile Gln Val Val Met Arg Gln Val Phe Phe	357	357
Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Ile Asp Pro Leu	358	358
Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr	359	359
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu	360	360
Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys	361	361
Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr	362	362
Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp	363	363
Leu Thr Leu Leu Arg Pro Ile His Ile His Lys Leu Leu Ser Ile	364	364
Leu Val Ala Leu Asn Ala Leu Thr Pro Lys Met Val Arg Arg Asn	365	365
Thr Leu Ile Arg Leu Met Val Lys Thr Leu Arg Ile Lys Ser Leu	366	366
Leu Gly Leu Pro Ala Ile Ser Tyr Ser Phe Ile Val Ile Arg Asn	367	367
Ala Lys Val Ala Ala Tyr Ile	368	368

[illegible]

Gln Glu Gln Leu Leu Leu Arg Gln His Tyr Gln Lys Lys Phe Lys	100	105	110
AAU AGU AAU TAU TUA AUA AUA TAT TAT GTA GAT GTG GTA TAU AAT TTT	115	120	125
Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Val Leu Tyr Thr Phe	130	135	140
GCA AAC TGC TCA GGA CTG GAG ITG ATC TTT GGG CTA AAT GGG TTA	145	150	155
Ala Asn Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu Asn Ala Leu	160	165	170
TTA AGA AUA GUA GAT TTA CAG TTA AAT AAT TAT AAT GTT CAG TTT	175	180	185
Leu Arg Thr Ala Asp Leu Gln Trp Asn Ser Ser Asn Ala Gln Leu	190	195	200
ATC CTG GAG TAC TAT TAT TCC AAG GGG TAT AAC ATT TAT TGG GAA	205	210	215
Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile Ser Trp Glu	220	225	230
CTA GGC AAT GAA GAT AAG AGT TTC CTT AAG AAG GGT GAT ATT TTC	235	240	245
Leu Gly Asn Glu Pro Asn Ser Phe Leu Lys Lys Ala Asp Ile Phe	250	255	260
ATC AAT GGG TCG CAG TTA GGA GAA GAT TAT ATT CAA TTG CAT AAA	265	270	275
Ile Asn Gly Ser Gln Leu Gly Glu Asp Tyr Ile Gln Leu His Lys	280	285	290
CTT CTA AGA AAG TCC ACC TTC AAA AAT GCA AAA CTC TAT GGT CCT	295	300	305
Leu Leu Arg Lys Ser Thr Phe Lys Asn Ala Lys Leu Tyr Gly Pro	310	315	320
GAT GTT GGT CAG GAT CGA AGA AAG ACG GGT AAG ATG CTG AAG AGC	325	330	335
Asp Val Gly Gln Pro Arg Arg Lys Thr Ala Lys Met Leu Lys Ser	340	345	350
TTT CTG AAG GGT GAT GGA GAA GTG ATT GAT TCA GTT ACA TAC CAT	355	360	365
Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Val Thr Trp His	370	375	380
TAU TAU TAU AUU AUA GUA UUA AUA UUA AUA AUA AUA AUA AUA AUA	385	390	395
His Tyr Tyr Leu Ala Gly Arg Ala Ala Thr Arg Gln Arg Phe Leu	400	405	410
AAU UUA AUA UUA UUA UUA UUA UUA UUA UUA UUA UUA UUA UUA	415	420	425
Asn Pro Asp Val Leu Asp Ile Phe Ile Ser Ser Val Gln Lys Val	430	435	440
TTT CAG GTT GAT GAT AUA AUA AUA AUA AUA AUA AUA AUA AUA AUA	445	450	455
Phe Gln Val Val Asp Ser Thr Arg Ile Arg Lys Lys Val Trp Leu	460	465	470
AAU AUA AUA AUA AUA AUA AUA AUA AUA AUA AUA AUA AUA AUA	475	480	485
Asn Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg Arg	490	495	500

TTA GGT GGA ATG GGA ATA GAA GTG GTG ATG AAG CAA GTA TTT TTT	1577
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Val Val Phe Phe	
440 441 442	
GGA GCA GGA AAC TAT CAT TTA GTG GAT GAA AAC TTC GAT GGT TTA	1443
Gly Ala Gly Asn Tyr His Leu Val Asp Phe Asn Phe Asp Phe Leu	
443 444 445	
GCT GAT TAT TTG CTA TCT GCT CTG TTC AAG AAA TTG GTC GGT AAT	1488
Phe Asp Tyr Thr Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr	
446 447 448	
AAG GTG TTA ATG GGA AAG GTG CAA GGT TGA AAG AGA AAG AAG GGT	1533
Lys Val Leu Met Ala Ser Val Gln Gly Ser Lys Arg Arg Lys Leu	
470 471 472	
GGA GTA TAC GGT CAT TTG ACA AAC AAT GAC AAT CAA AGG TAT AAA	1578
Arg Val Tyr Leu His Cys Thr Asn Thr Asp Asn Pro Arg Tyr Lys	
480 481 482	
GAA GGA GAT TTA ACT CTG TAT GCC ATA AAG CTC CAT AAG GTC ACC	1623
Glu Gly Asp Leu Thr Leu Tyr Ala Ile Asn Leu His Asn Val Thr	
500 501 502	
AAG TAC TTG CGG TTA CCC TAT GCT TTT TCT AAG AAG CAA GTG GAT	1668
Lys Tyr Leu Arg Leu Pro Tyr Pro Phe Ser Asn Lys Gln Val Asp	
515 516 517	
AAA TAC GGT CTA AGA CCT TTG GGA GGT CAT GGA TTA GGT TCC AAA	1713
Lys Tyr Leu Leu Arg Pro Leu Gly Pro His Gly Leu Leu Ser Lys	
530 531 532	
TCT GTC CAA CTC AAT GGT CTA ACT CTA AAG ATG GTG GAT CAT CAA	1758
Ser Val Gln Leu Asn Gly Leu Thr Leu Lys Met Val Asp Asp Glu	
545 546 547	
AAT TTG CAA GGT TTA ATG CAA AAA GGT GTC CTA CTA GGA AAT TGA	1803
Asn Phe Leu Leu Leu Met Arg Leu Phe Leu Arg Phe Gly Ser Phe	
550 551 552	

TTA GGT GGA ATG GGA ATA GAA GTG GTG ATG AAG CAA GTA TTT TTT 1577  
Ser Ala Arg Met Gly Ile Glu Val Val Met Arg Val Val Phe Phe  
440 441 442

GGA GCA GGA AAC TAT CAT TTA GTG GAT GAA AAC TTC GAT GGT TTA 1443  
Gly Ala Gly Asn Tyr His Leu Val Asp Phe Asn Phe Asp Phe Leu  
443 444 445

GCT GAT TAT TTG CTA TCT GCT CTG TTC AAG AAA TTG GTC GGT AAT 1488  
Phe Asp Tyr Thr Leu Ser Leu Leu Phe Lys Lys Leu Val Gly Thr  
446 447 448

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16

ATTACIATAT GGAAGCTTG GGGGAGGCG GGGCTCTTA TTCTTTAAI CAGGAATTGA 70  
TAAAGAATTI TGGTCTTG ATCTTTT AATCTCACT TAGCTATCT TCACTTAA 130  
TTTCTTACG CAAAGTAAA AATCTTAA AATCTCTT CAAAGTAA ATCAGATTTT 190  
GGTGGCT AATCTTAA AATCTTTAT AATCTTAA CAAAGTAA GATCTAATAI 250  
TCAATGGAG GTTTTATG AGGCTTACG CATACTT CAAAGTAA ATCTTATCT 310  
GGAGTCGGAA AGGCTGGTT CCAAGGAG CCAAGGAA AGGCTGTA GGAAGGCTGG 370  
TCCGGGATC CCAAGGCTG TCCGAGGAG CTCTCTCGG GGGCTGCTC CCAAGGCTC 430  
CCGGGGCTT GGATCTCGG CATCTCGGA CCTTAAGT GGGCTGCGT GATTCTGTAA 490  
GTGAACGTA CCAAGGCTG GGGGAGGAG AATCTTAA TAGGAGAG CCGGCTAGG 550  
GGGGGCTT TCAATCTA CCAAGGAG CATACTT AATCTTAA CCAAGTAA 610

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17

CCGAGGAG AGGAGATCA G 11

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 21  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18

AGGCTTCGAG CCGAGGAGCA T 11

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19

GAATATGAG TCAATATA G 12

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20

AATATACT AATCTCTT 14

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21

AATATACT AATCTCTT 14

## (1) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:  
 ATTCTATA ATCTCTATA CA 17

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:  
 GCATCTTAGC CTTCTTTT CC 23

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:  
 GAGGAGCCAG GTGAGCCCAA GAT 23

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:  
 TTGATCCCA AGAAGGAAT AA 23

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:  
 AATTATATA ATCTCTATA CA 23

## (2) INFORMATION FOR SEQ ID NO:27:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid

## (2) INFORMATION FOR SEQ ID NO:28:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23  
 (B) TYPE: nucleic acid

(A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10  
 (GATGTTA)G CATTCTTTT (3) 11

(ii) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 14  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11  
 (GTAGTGATGC CATGTAACG AATG) 24

(iii) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12  
 AGGCACCCTA GAGATGTTCC AG 22

(iv) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13  
 GAAGATTTCT GTTCATGA GGTG 24

(v) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14  
 (A AAT GA) AATAATG AATG 25

(vi) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 17  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15

(A) LENGTH: 17  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear



(1) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14  
 GATAGTGA AACTGTG A 11

(2) INFORMATION FOR SEQ ID NO:15:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15  
 AAATATGCT TATTAAGAC TT 11

(3) INFORMATION FOR SEQ ID NO:16:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16  
 GGCAGGCTGG CGTGGATGGT GA 11

(4) INFORMATION FOR SEQ ID NO:17:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17  
 GTGGATGGTG ATGGACAGGA AG 11

(5) INFORMATION FOR SEQ ID NO:18:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 23  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18  
 GTAATACGAC TACTATAGG GT 11

(6) INFORMATION FOR SEQ ID NO:19:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19  
 AATATATG AACTAT 11

SEQUENCE INFORMATION FOR SEQ ID NO:20:

(7) INFORMATION FOR SEQ ID NO:21:  
 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 11  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21  
 AATATATG AACTAT 11

1. INFORMATION FILE REF. ID: A-141

(1) PRIMER: CHABA TERPENTIN:

- (A) LENGTH: 44-48  
(B) TYPE: random  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(2) SEQUENCE DESCRIPTION: REF. ID: A-141

ATTGATTATA GGGCTGAGG GGGT...

2. INFORMATION FILE REF. ID: A-142

(1) PRIMER: CHABA TERPENTIN:

- (A) LENGTH: 44-48  
(B) TYPE: random  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(2) SEQUENCE DESCRIPTION: REF. ID: A-142

GGATTTTGGT TGAATGCAAT GTTGGCTTGC CATGCAATTT TTAATGATCA 150  
GGCTGTTGAT TAGCTTGCAAT TATGCTTTG GATGCAATTT GATGCTTACA 160  
GATGCTTGGC CAGGCTGCTC TTGAA TETT GAGCTTAGT GATGCAATTT 170  
GTTTGGCTTT CCAAAATGCT GGGATTAAG GGTGGAATTA TGAATGCTTT 200  
GGGCTGCTTT CCATATTAAT AATGCAATG TAAAC AGAA GGATGCACTA 250  
TTTAAAGAAA TTGCAATGCT GAGCTTTTCA AATGCAATTA ACATGTTAAA 300  
GAAATTTGCA TGAATGCTTT TGGCAATGCT GATGCAATTT GATGCTTAG 350  
CATTTTGTGA GGGCTGAGAG GGGCAATTA GAGGTTAGGA GATGCAATTT 400  
ATCTTGAGAG ACATGCTGAA ATGCTTCTC TACTAGAAA ACATGCAATTT 450  
TAGTGGGGAG TGAATGCTTT GGGCTTGAAT GATGCTTTA TGGGAGCTTT 500  
AGGCTGAGGA ATGCTTCTCA TGGGAGCTTT GATGCTTTA TGGGAGCTTT 550  
ATGCTGCTTT TGAATGCTTT GGGCTTGAAT GATGCTTTA TGGGAGCTTT 600  
AAAAAGAAA AAGGAGAAA ATGCTTGAAT GATGCTTTA TGGGAGCTTT 650  
TGAAGGGT GTGGCTGCTTT GATGCTTTA TGGGAGCTTT GATGCTTTA 700  
AGAACTTT GTTTTCTTCA CTGCTTGAAT GATGCTTTA TGGGAGCTTT 750  
TTCTGCTAAT GTGGCTTTTA TGTGCTAAT TTTTCTTCA TGTGCTAAT 800  
AATGCTTGA TGAATGCTTT ATGCTTGAAT TTTTCTTCA TGTGCTAAT 850  
GCTTTTGAAT TGTGCTTTA TGTGCTTGA TGTGCTTTA TGTGCTTGA 900  
TTATTTATTT TTTTGAATTT GGTGCTTGA TGTGCTTTA TGTGCTTGA 950  
AGTCTGCTTT TGAATGCTTT GGTGCTTGA TGTGCTTTA TGTGCTTGA 1000  
AGTCTGCTTT TGAATGCTTT GGTGCTTGA TGTGCTTTA TGTGCTTGA 1050  
AGTCTGCTTT AATTTTCTTT ATTTTCTTTA TGTGCTTTA TGTGCTTGA 1100  
TGTGCTTTA GTTCTTGAAT TGTGCTTTA TGTGCTTTA TGTGCTTTA 1150  
GCTGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA 1200  
AGCTGCTTTA GCTGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA 1250  
TATTTTATTT TGTGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA 1300  
AAGCTTTTGT GTGCTTTTGT CTTGCTTTA TGTGCTTTA TGTGCTTTA 1350  
CTAAGCAGCA ACATGCTTGA ATGCTTGAAT TGTGCTTTA TGTGCTTTA 1400  
AGCAAACTTT AATCTTCTTT TGTGCTTTA TGTGCTTTA TGTGCTTTA 1450  
GAAATGCTTT AATGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA 1500  
TGTGCTTTA AATGCTTTA TGTGCTTTA TGTGCTTTA TGTGCTTTA 1550

TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 160  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 170  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 180  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 190  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 200  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 210  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 220  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 230  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 240  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 250  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 260  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 270  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 280  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 290  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 300  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 310  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 320  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 330  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 340  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 350  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 360  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 370  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 380  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 390  
TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT TGAATGCTTT 400

[illegible]





[illegible][illegible]

PTTAAR AAG TGAACAGAGG AAGCGCTTT TTTTTTTTTT TTGAGATGGA 18800  
 AT TGA TGT TGTGCTGAG C TGAAGTGT AATGAGG AA TTTTCTTA 18810  
 CTGAA T CAGCTGCA GCTTAAGAG TTCTGCTT TTGCTTCT 18900  
 ATGATTTT AGGAATATTT GATTATTAG TTCTGCTT TGAAGTAT 18910  
 PAAAGGAG TATATATTATG ATTATGAG A TGAATTT AAGGCT 19000  
 GAGGAGGAG TGTGTAAT TTTGCTATTT TTTAGTA AG AATGCTTT 19010  
 AGCTGCTG TGGGCTA GCTGCTT GAGTCTT GAGTCTT 19100  
 PAAAGGAG TATGCTT TGAATTT TGAATTT TGAATTT 19110  
 NCTCTGCT TATGCTT TTTTATG TATGCTT TATGCTT 19200  
 TTTATTT TATGCTT TATGCTT TATGCTT TATGCTT 19300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 19900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20000  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20100  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20200  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 20900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21000  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21100  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21200  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 21900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22000

TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22100  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22200  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 22900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23000  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23100  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23200  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 23900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24000

TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24100  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24200  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24300  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24400  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24500  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24600  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24700  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24800  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 24900  
 TATGCTT TATGCTT TATGCTT TATGCTT TATGCTT 25000





CGCTAAAT	GCTGAGAG	TAGGAAAT	ATGATGAGA	ATGA TTAT	26800
TTTCTTCT	TTTG TTCT	AGAGA AGT	TCTTCTT T	AGTAACT	26850
GA TGGTGG	GTAAGATAT	GCTT A TG	AA TTGGG	TTTGGCT	26900
TAGAAAT	CTGAT TGAG	TGGA AGAT	AG TGGGCT	ATA TGAAT	26950
ATCAATCA	CTGG TACT	TAGAAAT	TTTCTTCT	AT TCTT	27000
CTCTTTTG	CTGA CTGA	TCTTTTAT	T TGGT T	TAGTATCT	27050
TCATCTTG	CTCT AGAG	TG TCTT	AT TGAAG	CTGAGA	27100
CACTCTCA	CTTTT TTA	AAAAAATA	AGTTCTT T	TGTGAGAA	27150
TCTTATAG	TG AATCT	ATTAAGAT	CT TCTT	TATATATAT	27200
TTTAAAGA	CTAT AATTT	TG TCTT	TATATAT	CTAT TATG	27250
CAATCTT	CTAT ATTA	AACTT TAT	TATATAT	CACTT TCT	27300
TTT TAT	CTATAT AT	AA TCTT	TTT TAT	ATTAAT T	27350
ATG CTAT	TTCT ATAT	ATTT TTTT	TTT TAT	ATTAATATA	27400
TTTTTTT	TGCTT TG	AA TCTT	CT TATAT	ATTTATCT	27450
CTATCT	AT TATATG	TT TAT	TTTATAT	AAAAATA	27500
TAT TTTAT	TTATAT TA	ATTT TAT	TTTATAT	TTTAATAT	27550
ATTA TAT	ATTT AATA	AAAAAAT	TTTCTT	TAAT TCT	27600
CTT ATTT	ATTTATAT	AAAAAAT	TTTATAT	TTATATAT	27650
TATCTT	TTATATCT	TTTTATA	AT TAAAT	TTATATAT	27700
CTATGAAT	TTATATCT	TTTTATA	AT TAAAT	TTATATAT	27750
CTCTAT	ATTTCTT	TTTTATA	AT TAAAT	TTATATAT	27800
ATTTCTAT	TTTCTAT	TTTTATA	AT TAAAT	TTATATAT	27850
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	27900
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	27950
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28000
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28050
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28100
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28150
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28200
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28250
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28300
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28350
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28400
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28450
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28500
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28550
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28600
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28650
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28700
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28750
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28800
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28850
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28900
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	28950
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29000
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29050
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29100
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29150
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29200
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29250
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29300
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29350
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29400
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29450
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29500
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29550
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29600
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29650
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29700
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29750
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29800
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29850
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29900
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	29950
TTTCTAT	TTTTATA	TTTTATA	AT TAAAT	TTATATAT	30000

ATAGGCTAT	TATCATTTTT	TTTTAGCT	CAGAGTTAA	A TGAATTTT	9880
TTTCAAAA	TAGTTAAT	TTCTGCTG	AAATAGAAAT	GAGGTCTTAA	9885
AGTTTAAAG	CAGATGAAAT	TAGGAGAT	CTCTCTCT	TCTCTAT	9890
ATTTTCTA	GTATATAATT	TTTAAAAT	TTTTTCTG	CTATCTCT	9895
AGATTTAT	GTAGGAGAA	TTCTGCTG	CTAGAAAT	TTCTTTAG	9900
AGATTTAT	TAGGATAT	TAATGAGT	CTCTCTA	AATGAGAT	9905
CTCTTAA	ATGAATAAAT	ATGAATAA	ATGATAT	ATGAATAA	9910
AGATTTAT	GAATTTAA	TTCTGCTG	TTCTTTAA	AATAGAA	9915
ATTTTCTA	TCTCTCT	CTCTCT	CATTAATTT	TTCTAGT	9920
CTAGGCTA	ATAGGATTT	TTGATAT	CTCTTT	CTAATAGCA	9925
CTATTTAT	ATGATTAAT	ATGATTT	CTCTTT	ATATGATAT	9930
TTCTTTAA	ATATTTAA	ATGATTT	CTCTTT	CTCTTT	9935
CTTAACTA	CTATTAAT	CTCTTT	CTCTTT	CTCTTT	9940
CTAGGCTA	TTATAGGAT	CTCTTT	CTCTTT	CTCTTT	9945
CTCTTTAA	ATAGGAAAT	CTCTTT	CTCTTT	CTCTTT	9950
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9955
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9960
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9965
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9970
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9975
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9980
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9985
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9990
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	9995
CTCTTTAA	CTATTTAA	CTCTTT	CTCTTT	CTCTTT	10000





```

(19) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 2396
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: double
      (D) TOPOLOGY: linear
(20) SEQUENCE DESCRIPTION: SEQ ID NO:48

```

```

AGTCTGAG 106 TTCTT 11 TGGTGTG AGGAAATCG AGTGGAAATG 1600
AGTCTGAG 114 TTCTT 12 TGGTGTG AGGAAATCG AGTGGAAATG 1650
AGGACAG 121 CTACTGAG 13 AGTGGAAATG AGTGGAAATG 1700
AAGTGTG 131 GTTAAAG 14 GATCTTTGGT CTAAATGAGT TACTAAAG 1750
CTAGGTTA GTTAAAG 15 GATCTTTGGT CTAAATGAGT TACTAAAG 1800
CTCTTCCAA GGTATATAAG ATCTG 166 TAAAGGCAAG TAAAGGCAAG 1850
AGTTTTCGA AAAAAAG 17 AATTCTT AT GATGGGTTG AATTAGAGA 1900
AGACTTGTG TGTTCCTA AACTT TACA AAGTGTGCT TT TAAAAATG 1950
AAAACTCTA TGTTCCTA AATTCTT AT GATGGGTTG AATTAGAGA 2000
TGTTCCTA TGTTCCTA AATTCTT AT GATGGGTTG AATTAGAGA 2050
ATAGCATTA TATTATTA ATGAA 18 TAAAGGCAAG TAAAGGCAAG 2100
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2150
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2200
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2250
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2300
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2350
TAAAGGCAAG TAAAGGCAAG AATTCTT AT GATGGGTTG AATTAGAGA 2396

```

## (C) INFORMATION FOR SEQ ID NO:44:

## (i) SEQUENCE CHARACTERISTICS:

```

(A) LENGTH: 535
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

```

## (ii) SEQUENCE DESCRIPTION: SEQ ID NO:44

```

Met Leu Arg Leu Leu Leu Leu Trp Leu Trp Gly Pro Leu Gly Ala
1 10 15
Leu Ala Gly Gly Ala Leu Ala Gly Thr Ala Leu Thr Arg Arg Val
16 25 30
Val Arg Leu Leu Leu Trp Thr Leu Arg Leu Leu Arg Leu Ala
31 40 45
Leu Arg Leu Leu Leu Thr Thr Leu Thr Ala Leu Leu Leu Leu Arg
46 55 60
Pro Arg Thr Leu Thr Thr Leu Gly Ser Pro Arg Leu Arg Ala Leu
61 70 75
Ala Arg Gly Leu Ser Pro Ala Tyr Leu Arg Thr Gly Gly Thr Lys
76 85 90
Thr Arg Thr Leu Thr Thr Arg Leu Arg Leu Thr Thr Thr Thr
91 100 105

```

Gln Phe Lys Asn Ser Thr Tyr Ser Ala Ser Ser Val Asp Met Leu	118	119	120
Tyr Ser Ile Ala Lys Cys Ser Gly Leu Asp Leu Ile Phe Gly Leu	121	122	123
Asn Ala Leu Leu Arg Thr Phe Asp Leu Arg Ile Asn Ser Ser Asp	124	125	126
Ala Gln Leu Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile	127	128	129
Ser Trp Glu Leu Gly Asn Glu Pro Arg Ser Phe Trp Lys Lys Ala	130	131	132
His Ile Leu Ile Asp Gly Leu Gln Leu Gly Gln Asp Phe Val Glu	133	134	135
Leu His Lys Leu Leu Gln Arg Ser Ala Phe Gln Asn Ala Lys Leu	136	137	138
Tyr Gly Pro Asp Ile Gly Gln Pro Arg Gly Lys Thr Val Lys Leu	139	140	141
Leu Arg Ser Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Leu	142	143	144
Thr Trp His His Tyr Tyr Leu Asn Gly Arg Ile Ala Thr Lys Glu	145	146	147
Asp Phe Leu Ser Ser Asp Ala Leu Asp Thr Phe Ile Leu Ser Val	148	149	150
Gln Lys Ile Leu Lys Val Thr Lys Glu Ile Thr Pro Gly Lys Lys	151	152	153
Val Trp Leu Gly Glu Thr Ser Ser Ala Tyr Gly Gly Gly Ala Pro	154	155	156
Leu Leu Ser Asn Thr Phe Ala Ala Gly Phe Met Trp Leu Asp Lys	157	158	159
Leu Gly Leu Ser Ala Gln Met Gly Ile Glu Val Val Met Arg Gln	160	161	162
Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Phe	163	164	165
Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Phe Lys Lys Leu	166	167	168
Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg	169	170	171
Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro	172	173	174
Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His	175	176	177
Asn Val Thr Ser His Leu Lys Val His Phe Thr Ser Phe Arg Lys	178	179	180
His Val Arg Leu Thr Leu Lys Tyr His Arg Gly Ile Asp Gly Leu	181	182	183
Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val	184	185	186
Asp Glu Gln Thr Leu His Ala Leu Thr Gln Lys Pro Leu Pro Ala	187	188	189
Arg Tyr Ala Leu Ser Leu His Ala Ile Ser Trp Gly Ile His Val	190	191	192

															TTT TTT AGT		
TTT	TTT	TAA	GAA	ATG	TGG	GAT	CAG	GTT	TTT	GAA	GGG	A	A	AAG	AGA	50	
TAT	TAA	GAT	ATT	GAA	CTG	AGG	AAA	T	TAT	TTG	GGT	CAG	ATT	TTT	55		
GAA	GGG	TAA	AGT	GAT	GGG	TGT	CTT	AAG	GAA	GAA	GGG	AGT	TTG	ATT	60		
AAG	GGT	TGG	CTG	AAG	CTT	CAA	G	G	GGG	ATT	TAA	AGT	CTG	TTT	65		
GTT	CTG	GGG	GAA	TAT	GAA	TGG	GAT	CTA	GAA	AAG	GTG	A	A	GAG	70		
GAA	GAA	CTG	CAA	AAA	G	T	GGT	GGT	A	G	CTT	GGG	GGG	TTT	75		
GCT	GGG	TGG	GAA	GGT	CTG	CGA	TAA	T	T	TTG	ATT	CTG	GAT	A	80		
GAA	GGG	TTT	GAA	TGG	CTG	TGG	ATG	ATT	TGA	AAA	GTG	GAT	GTG	AG	85		
GGG	GGG	GAG	GGG	AAA	G	AGT	AAG	GAA	AGA	GAG	AGG	GAG	GAG	90			
GGG	AGG	GGG	GGG	AGG	GAA	GGG	GGG	TAG	GGA	GGG	AGT	GGG	GGG	AGG	95		
GGT	GGG	AGG	GAT	CGA	G	G	CTG	TGG	GAG	GGT	AGT	GAG	TGG	GGG	100		
CAG	AGG	CGA	AGG	AGG	AGC	GGT	TGG	AGG	GGG	CTT	GAG	GAA	GGG	GGG	105		
GGG	CTG	GGG	CAG	CTC	TGG	GGG	CAG	GGG	GGG	GTG	CAG	GAA	GGT	GGG	110		
ATG	CTG	AGG	CTG	CTG	CTG	CTG	TGG	CTC	TGG	GGG	GGG	CTC	GGT	GGC	115		
Met	Leu	Arg	Leu	Leu	Leu	Leu	Trp	Leu	Trp	Gly	Pro	Leu	Gly	Ala			
															5	10	15
CTG	GGC	CAG	GGG	GGC	GGC	GGG	GGG	ACC	GGG	GGG	ACC	GAC	GAC	GTG	120		
Leu	Ala	Gln	Gly	Ala	Pro	Ala	Gly	Thr	Ala	Pro	Thr	Asp	Asp	Val	20	25	30
CTA	GAC	TTG	GAG	TTT	TAC	ACC	AAG	GGG	GGG	CTC	CGA	AGG	GTG	AGT	35	40	45
Val	Asp	Leu	Glu	Phe	Tyr	Thr	Lys	Arg	Pro	Leu	Arg	Ser	Val	Ser			
GGG	TGG	TTG	CTG	TGG	ATC	AGT	ATC	GAC	GGG	ACC	CTG	GGG	ACC	GAC	50	55	60
Pro	Ser	Phe	Leu	Ser	Ile	Thr	Ile	Asp	Ala	Ser	Leu	Ala	Thr	Asp			
GGG	GGG	TTG	CTC	ACC	TTG	CTG	GGC	TAT	GGA	AGG	CTC	GGT	GCT	CTG	65	70	75
Pro	Arg	Phe	Leu	Thr	Phe	Leu	Gly	Ser	Pro	Arg	Leu	Arg	Ala	Leu			
GCT	AGA	GGG	TTA	TGT	GGT	GGA	TAC	TTG	AGA	TIT	GGG	GGG	AGA	AAG	80		
Ala	Arg	Gly	Leu	Phe	Phe	Ala	Tyr	Ile	Arg	Ile	Gly	Gly	Thr	Lys			

[illegible]



GAG TTT AAG AAT AGG AAT TAC TTA AAA AGG TTA GTG GAG ATG CTG 1377  
Gln Ile Lys Asn Ser Thr Tyr Ser Arg Ser Ser Val Asp Met Leu  
177 187 197

TAC AAT TTT GCG AAG TCG TCG GCG TTA GAG GTG ATG TTT GGT CTA 1132  
Tyr Ser Ile Ala Lys Lys Ser Gly Leu Asp Leu Ile Ile Gly Leu  
178 179 180

AAT GCG TTA CTA CGA AAT GGA GAG TTA GCG TGG AAG AGG TCG AAG 1178  
Asn Ala Leu Leu Arg Thr Pro Asp Leu Arg Trp Asn Ser Ser Asn  
181 182 183

GCG CAG CTT CTC CTT GAG TAC TCG TTT TCG AAG GGT TAT AAT ATG 1273  
Ala Gln Leu Leu Leu Asp Tyr Cys Ser Ser Lys Gly Tyr Asn Ile  
190 195 210

TCC TGG GAA CTG GGC AAT GAG CGC AAC AGT TTC TGG AAG AAA GCT 1208  
Ser Trp Glu Leu Gly Asn Glu Pro Asn Ser Phe Trp Lys Lys Ala  
215 220 225

GAG ATT CTC ATC GAT GGG TTG CAG TTA GGA CAA GAC TTT GTG GAG 1313  
His Ile Leu Ile Asp Gly Leu Gln Leu Gly Glu Asp Phe Val Glu  
230 235 240

TTG CAT AAA CTT CTA CAA AGG TCA GGT TTC CAA AAT GCA AAA CTC 1358  
Leu His Lys Leu Leu Gln Arg Ser Ala Phe Gln Asn Ala Lys Leu  
245 250 255

TAT GGT GGT GAC ATC GGT CAG GGT CGA GGG AAG ACA GTT AAA CTG 1403  
Tyr Gly Pro Asp Ile Gly Gln Pro Arg Gly Lys Thr Val Lys Leu  
260 265 270

CTG AGG AGT TTG CTG AAG GGT GGC GGA GAA GTG ATC GAC TGT CTT 1448  
Leu Arg Ser Phe Leu Lys Ala Gly Gly Glu Val Ile Asp Ser Leu  
275 280 285

ACA TGG CAT CAG TAT TAC TTG AAT GGA GGC ATC GGT AGG AAA GAA 1493  
Thr Trp His His Tyr Tyr Leu Asn Gly Arg Ile Ala Thr Lys Glu  
290 295 300

AT TTT TTA AAT TTT AAT TTT TTA AAT AAT TTT AAT TTT TTT 1538  
Arg Ile Leu Ser Ser Arg Arg Ser Arg Thr Ile Ile Leu Ser Val  
305 310 315

CAA AAA ATT CTG AAG GTC ACT AAA GAG ATC ACA GTT GGC AAG AAG 1583  
Gln Lys Ile Leu Lys Val Thr Lys Glu Ile Thr Pro Gly Lys Lys  
320 325 330

CTT TAT TTT GAT GAT AAT AAT TTA TTT TAT TTT AAT TTT TTT 1628  
Leu Tyr TTT GAT GAT AAT AAT TTA TTT TAT TTT AAT TTT TTT  
335 340 345

TTG GGG CTG TCA GGC CAG ATG GGT ATA GAA CTC GTC ATG AGG CAG 1718  
 Leu Gly Leu Ser Ala Gln Met Gly Ile Gln Val Val Met Arg Gln  
 375 376 377

GTC TTC TGC GGA GGA GGC AAC TAC GAT TTA GTC GAT GAA AAT TTT 1723  
 Val Phe Phe Gly Ala Gly Asn Tyr His Leu Val Asp Glu Asn Ile  
 378 379 380

GAG COT TTA CCT GAT TAC TGG CTC TCT CTT CTG TTC AAG AAA CTG 1808  
 Glu Pro Leu Pro Asp Tyr Trp Leu Ser Leu Leu Ile Lys Lys Leu  
 395 400 405

GTA GGT GGC AGG GTG TTA GTC TTA AGA GTC AAA GGT GGA GGC AGG 1853  
 Val Gly Pro Arg Val Leu Leu Ser Arg Val Lys Gly Pro Asp Arg  
 410 415 420

AGC AAA CTC CGA GTG TAT CTC CAC TGC ACT AAC GTC TAT CAC CGA 1898  
 Ser Lys Leu Arg Val Tyr Leu His Cys Thr Asn Val Tyr His Pro  
 425 430 435

CGA TAT CAG GAA GGA GAT CTA ACT CTG TAT GTC CTG AAC CTC CAT 1943  
 Arg Tyr Gln Glu Gly Asp Leu Thr Leu Tyr Val Leu Asn Leu His  
 440 445 450

AAT GTC ACC AAG CAC TTG AAG GTA CCG COT CCG TTG TTC AGG AAA 1988  
 Asn Val Thr Lys His Leu Lys Val Pro Pro Pro Leu Phe Arg Lys  
 455 460 465

CGA GTG GAT ACG TAC CTT CTG AAG COT TCG GGG GCG GAT GGA TTA 2033  
 Pro Val Asp Thr Tyr Leu Leu Lys Pro Ser Gly Pro Asp Gly Leu  
 470 475 480

CTT TCC AAA TCT GTC CAA CTG AAC GGT GAA ATT CTG AAG ATG GTG 2078  
 Leu Ser Lys Ser Val Gln Leu Asn Gly Gln Ile Leu Lys Met Val  
 485 490 495

GAT GAG CAG ACC CTG GCA GGT TTG ACA GAA AAA COT CTC CCG GCA 2123  
 Asp Glu Gln Thr Leu Pro Ala Leu Thr Glu Lys Pro Leu Pro Ala  
 500 505 510

GAA AAT GGA TTA AAT CTC CTT GGT TTT TAT GAT TTT TTT TTT 2168  
 Gly Asn Ala Leu Ser Leu Ile Ala Ile Ser Lys Gly Ile Ile Val  
 515 520 525

ATA AGA AAT GCC AAA ATC CTT GCT TGT ATA TGA AAA TAA AAG GGA 2213  
 Ile Arg Asn Ala Lys Ile Ala Ala Cys Ile  
 530 535

TAT GAT AAT CTT GAT ATA AAA TTT GAT TTT TTT TTT ATT GAT AAA 2258  
 Tyr Asp Asn CTT GAT ATA AAA TTT GAT TTT TTT TTT ATT GAT AAA  
 540 545

(K) TYPE: nucleic acid  
 (L) STRANDEDNESS: double  
 (M) TOPOLOGY: linear

(XII) SEQUENCE DESCRIPTION: SEQ ID NO:46

```

CTTCTTATG TGTGTTATG CTCTCTGCTT CTTCTCTT CTTCTTATGAA 50
ATTACTGAG GCTGGGAGAG CTCTGAGTAA AGAGCTGCTT GACTTCAGCT 100
TTTAGAGTAA GAGGCTATTG CAAGAGTTGA CTCTCTGCTT CTTCTCTGAG 150
ATTATGAGG CAGGCTTCTT CAGTACTCT CTTCTCTCA CTTCTCTGAG 200
CTCTCTAGG CTCTGAGGCTT TCTTAGAG CTTATCTCTT GCTGCTTGA 250
GATTCTGAG CAGCAAGACT GATTCTTA TTTTGTATG CAACTAGAG 300
CTTACTCTG AACTAGAG TTAGTATA TCTAGAGAA AATATATAT 350
TTCTGCTCT GAGGCTCTT CAGCTGCTT GTGA 385

```

(I) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 541  
 (E) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47

```

AAATCAGGAC ATATCTTGA CTTATTGGC TCTTGGTCTAT ATTGGAGGCA 50
TTTGTATTGA TTTTAAATAA CTTCTAAAAT ACTGCTATGA AAGTGTAA 100
CGTCATTTGC CACATGCTGC CATTAAGTGT CACCACCTGC AGTGGTCTAC 150
TTAGAGAACA CCGCAATGGA TGITAACACT GAAGGCGCTG CCGCGGCTC 200
CCGAGGCTCT GGATCCAGCG TTGAAGCTTG CCGGCGCTC CCGAGGCTCT 250
GGATCCAGCA CTGAGGCTATG CCGCGGCTC CCGAGGCTCT GGAGCTTGCT 300
AAGGAGTGG CTGCTAGCT CTGGGCTTTT GCTTTATTCT TATCAATGAC 350
AGCCCTGACC GCTTTCTCT CAGGGGTACT GTAATGCTT TTATTTCTAT 400
ATACAAGCTG CGATTTTGGC ATTTCTTATG AAAAAAACC CATAGGAAAA 450
GGCGGGCAGG CTTAGTGAGT TTCTGAGG GAGAGTTTT TCTGTTAGAG 500
CTGGCANGGT CTGCTCATCG ACCATCTTCA GGCCTCGTGC C 541

```